UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/487,191	01/19/2000	Rakesh Agrawal	AM9-99-0226	2881

7590   07/15/2002

John L Rogitz
Rogitz & Associates
Suite 3120
750 B Street
San Diego, CA 92101

EXAMINER
FLEURANTIN, JEAN B

ART UNIT	PAPER NUMBER
2172	

DATE MAILED: 07/15/2002

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 07-01)

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) ☐ Responsive to communication(s) filed on *23 April 2002* .

2a) ☒ This action is **FINAL**.   2b) ☐ This action is non-final.

3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) ☒ Claim(s) *14-19 are canceled. And 1-13 and 20-23* is/are pending in the application.

   4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5) ☐ Claim(s) _____ is/are allowed.

6) ☒ Claim(s) *1-13 and 20-23* is/are rejected.

7) ☐ Claim(s) _____ is/are objected to.

8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) ☐ The specification is objected to by the Examiner.

10) ☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

   Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

11) ☐ The proposed drawing correction filed on _____ is: a)☐ approved b)☐ disapproved by the Examiner.

   If approved, corrected drawings are required in reply to this Office action.

12) ☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

13) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

   a)☐ All b)☐ Some * c)☐ None of:

   1. ☐ Certified copies of the priority documents have been received.

   2. ☐ Certified copies of the priority documents have been received in Application No. _____ .

   3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

   * See the attached detailed Office action for a list of the certified copies not received.

14) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).

   a) ☐ The translation of the foreign language provisional application has been received.

15) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☐ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____ .

4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .

5) ☐ Notice of Informal Patent Application (PTO-152)

6) ☐ Other: .

DETAILED ACTION

1.   Claims 14-19 are canceled.

Claims 1-13 and 20-23 are remained for examination.

Response to Amendment

2.   Applicant's arguments submitted on 04/23/2002 with respect to claims 1, 7 and 13 have

been fully considered but are not persuasive.

Response to Applicant' remarks

3.   As per claims 1-13 and 20-23, Applicant argues that the Fayyad reference and the

Tendick reference together or individually do not teach or fairly suggest:

Applicant on pages 5 and 6, stated that 'Fayyad does not use perturbed values of original

values at all'. However, Examiner disagrees because Fayyad includes records that make up each

of the subsets any of the clusters have zero membership a new cluster mean for the empty cluster

is chosen by picking the mean of the entire data set and perturbing that mean by a small random

amount corresponding to a variance in the data mean in each dimension of the data sample;

which is read as perturbed values of original values (see col. 11, lines 14-19). Also, columns 1,

lines 40 through 52 Fayyad teaches there are various approaches to performing the optimization

problem of finding a good set of parameters the most effective class of methods is known as the

iterative refinement approach, the basic algorithm goes as follows initialize the model parameters

producing a current model, decide memberships of the data items to clusters assuming that the

current model is correct, re-estimate the parameters of the current model assuming that the data

memberships obtained in two are correct producing a new model, if the current model and new

model are sufficiently close to each other determine else go to two. And, also in column 2, lines

35 through 36, Fayyad further teaches the steps of the situations initialization is done by

randomly picking a set of starting points from the range of the data. Thus, it would have been

obvious to a person of ordinary skill in the art at the time the invention was made to modify the

teachings of Fayyad with the step of perturbed values of original. This modification would allow

the teachings of Fayyad to provide a statistical data analysis (see col. 1, line 14).

Applicant on page 7, stated that 'Fayyad does not directed to maintaining privacy'.

However, Examiner disagrees because Tendick includes the steps of a user of such an employee

database should be interested in statistical analysis alone and thus should have no legitimate

interest in such items as the salary of any individual employee, thus the DBMS must include

mechanism which allow statistical analysis but not access to data on individual database records;

which is read as maintaining privacy (see page 48, lines 3-7). Also, in page 60, lines 34 through

38, Tendick further teaches the steps of suppose this applicant has a record in our database and to

preserve the privacy of individual records 'say against illicit view by a finance company' we

apply random data perturbation answering queries on A or A* instead of on A. Thus, it would

have been obvious to a person of ordinary skill in the art at the time the invention was made to

modify the teachings of Fayyad and Tendick with the step of maintaining privacy. This

modification would allow the teachings of Fayyad and Tendick to provide optimal security (see

page 60, line 28).

Examiner is entitled to give claim limitations their broadest reasonable interpretation in

light of the specification.

Interpretation of Claims-Broadest Reasonable Interpretation

During patent examination, the pending claims must be 'given the broadest reasonable

interpretation consistent with the specification.' Applicant always has the opportunity to amend

the claims during prosecussion and broad interpretation by the examiner reduces the possibility

that the claim, once issued, will be interpreted more broadly than is justified. In re Prater, 162

USPQ 541,550-51 (CCPA 1969).

Claim Rejections - 35 U.S.C. § 103

4.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1-13 and 20-23 are rejected under 35 U.S.C. 103(a) as being unpatentable over

Fayyad et al. (US Pat. No. 6,115,708) in view of Tendick et al. 'A Modified Random

Perturbation Method for Database Security - 03/1994' ("Fayyad"), ("Tendick").

As per claim 1, Fayyad substantially teaches the steps of perturbing original data

associated with the user computer to render perturbed data (thus, some methods take the mean of

the global data set and perturb it K times to get the K initial means or simply pick K random

points from the data set, in most situations initialization is done by randomly picking a set of

starting points from the range of the data; which is readable as perturbing original data associated

with the user computer to render perturbed data) (see col. 2, lines 32-36);

using a distribution of the perturbed data, generating at least on estimate of a distribution

of the original data; (thus, the variance in result illustrated by these depictions is fairly common

even in low dimensions using data from well-separated Gaussians, these figures also illustrate the

importance of the problem of having a good initial or starting point each of the two data clusters

depicted in FIGS. 5A and 5B depict clustering from 2 different samples of the same size that

were obtained from the same database; which is readable as using a distribution of the perturbed

data, generating at least on estimate of a distribution of the original data) (see col. 6, lines 25-32);

and

using the estimate of the distribution of the original data, generating at least one data

mining model (thus, the end of the clustering process any of the clusters have zero membership

then the corresponding initial guess at this cluster centroid is set to the data point farthest from its

assigned cluster center, this procedure decreases the likelihood of having empty clusters after

reclustering from the "new" initial point, resetting the empty centroids to another point may be

done in a variety of ways; which is readable as using the estimate of the distribution of the

original data, generating at least one data mining model) (see col. 7, lines 32-39). But, Fayyad

does not explicitly indicate the step of the maintaining the privacy of a user of the computer as

claimed in the preamble. However, Tendick implicitly indicates the step of the DBMS must

include mechanisms which allow statistical analysis but not access to data individual database

records, which is readable as maintaining the privacy of a user of the computer (see page 48,

lines 6-7). Thus, it would have been obvious to a person of ordinary skill in the art at the time

the invention was made to modify the teachings of Fayyad and Tendick with the step of

maintaining the privacy of a user of the computer. This modification would allow the teachings

of Fayyad and Tendick to improve the accuracy and the reliability of the system and architecture

for privacy preserving data mining, and provide the optimal protection against such a problem

among all possible covariance structures (see page 61, lines 4-5).

As per claims 2 and 8, Fayyad substantially teaches a method as claimed, wherein

perturbed data is generated from plural original data associated with respective plural user

computers (thus, the end of the clustering process any of the clusters have zero membership then

the corresponding initial guess at this cluster centroid is set to the data point farthest from its

assigned cluster center, this procedure decreases the likelihood of having empty clusters after

reclustering from the "new" initial point, resetting the empty centroids to another point may be

done in a variety of ways; which is readable as wherein perturbed data is generated from plural

original data associated with respective plural user computers) (see col. 7, lines 32-39).

As per claims 3 and 9, Fayyad wherein the original values cannot be reconstructed from

the respective perturbed values (thus, if at the end of the clustering process any of the clusters

have zero membership then the corresponding initial guess at this cluster centroid is set to the

data point farthest from its assigned cluster center this procedure decreases the likehood of

having empty clusters after reclustering from the new initial point resetting the centroids to

another point may be done in a variety of ways, which is readable as wherein the original values

cannot be reconstructed from the respective perturbed values) (see col. 7, lines 37-39).

As per claims 4 and 10, Fayyad substantially teaches a method as claimed, wherein at

least some of the data is perturbed using a uniform probability distribution (thus, the result of

clustering two different subsamples drawn from the same distribution and initialized with the

same starting point, which is readable as wherein at least some of the data is perturbed using a

uniform probability distribution) (see col. 6, lines 22-26).

As per claims 5 and 11, Fayyad substantially teaches a method as claimed, wherein at

least some of the data is perturbed using a Gaussian probability (thus, the model cluster is

assumed be a Gaussian for each cluster the Gaussian is centered at the mean of the cluster, which

is readable as wherein at least some of the data is perturbed using a Gaussian probability) (see

cols. 2-3, lines 67-2).

As per claim 6, Fayyad substantially teaches a method as claimed, wherein at least some

of the data is perturbed by selectively replacing the data with other values based on a probability

(thus, a multinomial distribution has a simple set of parameters for every attribute a vector of

probabilities specified the probabilities of each value of the attribute given the cluster, which is

readable as wherein at least some of the data is perturbed by selectively replacing the data with

other values based on a probability) (see col. 10, lines 20-25).

As per claims 7 and 13, in addition to the discussion in claim 1 above, Fayyad further

teaches the steps of sending the perturbed values to a server computer not having access to the

original values (thus, data samples are each chosen as a starting point for a clustering of all the

candidate solutions , the best solution returned as the refined 'improved' starting point to be used

in clustering the full data set; which is readable as sending the perturbed values to a server

computer not having access to the original values) (see col. 3, lines 37-41). Also, in column 11,

lines 36 through 40, Fayyad further teaches steps of finding multiple candidate clustering starting

points from the multiple data subsets retrieved from the database and choosing an optimum

solution from the multiple number of candidate clustering starting points to begin subsequent

clustering on data in the database.

As per claim 12, Fayyad substantially teaches a method as claimed, wherein the method

acts further comprise perturbing categorical values of at least some categorical attributes by

selectively replacing the categorical values with other values based on a probability (thus, assume

the user is clustering using the EM algorithm and that data is discrete, and hence each cluster

specifies a multinomial distribution over the data a multinomial distribution has a simple set of

parameters for every attribute a vector of probabilities specified the probabilities of each value of

the attribute given the cluster, since these probabilities are continuous quantities they have a

"centroid" and K-means can be applied to them; which is readable as wherein the method acts

further comprise perturbing categorical values of at least some categorical attributes by

selectively replacing the categorical values with other values based on a probability) (see col. 10,

lines 18-25).

As per claims 20 and 23, Fayyad substantially teaches a method as claimed, further

comprises sending the model to at least one user computer for use thereof by the user computer

on original data (thus, a mixture model M having K clusters Ci, I=1 . . . , K assigns a probability

to a data point x as follows ##EQU1## where W.sub.i are called the mixture weights, the

problem of clustering is identifying the properties of the clusters Ci. Usually it is assumed that

the number of clusters K is known and the problem is to find the best parameterization of each

cluster model; which is readable as sending the model to at least one user computer for use

thereof by the user computer on original data) (see col. 1, lines 25-38).

As per claims 21, Fayyad substantially teaches a method as claimed, wherein the user

computer uses the model on original data to render a classification, and then sends the

classification to the Web site (thus, each of the points in figure 4B may be thought of as a "guess"

for the possible location of a mode in the underlying distribution the estimates are fairly varied

but they exhibit "expected" behavior the subsampling produces a good separation between the

two clusters; which is readable wherein the user computer uses the model on original data to

render a classification, and then sends the classification to the Web site) (see col. 6, lines 13-15).

As per claim 22, Fayyad substantially teaches a method as claimed, wherein the model is

sent to the user computer as a JAVA applet (see col. 1, lines 20-35).

5.   The prior art made of record and not relied upon is considered pertinent to applicant's

disclosure. Fayyad et al. US patent Number 6,012,058 relates to database analysis. Agrawal et

al. US patent Number 6,233,575 relates to organizing and indexing information items. Fayyad et

al. US patent Number 6,263,337 relates to data sets that characterize the data.

6.   Applicant's amendment necessitated the new ground(s) of rejection presented in this

Office action. Accordingly, **THIS ACTION IS MADE FINAL.** See MPEP § 706.07(a).

Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE

MONTHS from the mailing date of this action. In the event a first reply is filed within TWO

MONTHS of the mailing date of this final action and the advisory action is not mailed until after

the end of the THREE-MONTH shortened statutory period, then the shortened statutory period

will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR

1.136(a) will be calculated from the mailing date of the advisory action. In no event, however,

will the statutory period for reply expire later than SIX MONTHS from the date of this final

action.

Conclusion

7.   Any inquiry concerning this communication from examiner should be directed to Jean

Bolte Fleurantin at (703) 308-6718. The examiner can normally be reached on Monday through

Friday from 7:30 A.M. to 6:00 P.M.

If any attempt to reach the examiner by telephone is unsuccessful, the examiner's supervisor, Mrs. KIM VU can be reached at **(703) 305-8449.** The FAX phone numbers for the Group 2100 Customer Service Center are: *After Final* **(703) 746-7238,** *Official* **(703) 746-7239,** and *Non-Official (703) 746-7240.* NOTE: Documents transmitted by facsimile will be entered as official documents on the file wrapper unless clearly marked *"DRAFT"*.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the Group 2100 Customer Service Center receptionist whose telephone numbers are **(703) 306-5631, (703) 306-5632, (703) 306-5633.**

Jean Bolte Fleurantin

July 11, 2002

JBF/

KIM VU
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100